ON SELECTING A SUBSET CONTAINING THE POPULATION WITH THE SMALLEST VARIANCE [1/]

Shanti S. Gupta
Bell Telephone Laboratories

Milton Sobel
University of Minnesota

Special Report No. 4

Department of Statistics
University of Minnesota
Minneapolis, Minnesota

[1/] **This work was started and the tables were constructed at the Bell Telephone Laboratories.**

ON SELECTING A SUBSET CONTAINING THE POPULATION WITH THE SMALLEST VARIANCE

by

Shanti S. Gupta and Milton Sobel*
Bell Telephone Laboratories

1. Summary

A multiple decision approach is taken to the problem of selecting a subset from k given normal populations which includes the "best" population, i.e., the one with the smallest population variance. The population variances of the k normal populations are unknown and the population means may be known or unknown. Based on a common number of observations from each population, a procedure R is defined which selects a subset which is never empty, small in size and yet large enough to guarantee with preassigned probability that it includes the best population, regardless of what are the true unknown population variances. Expressions for the probability of a correct selection using R are derived and it is shown that, for the case in which the k sample variances have a common number, ν, of degrees of freedom, the infimum of this probability is identical with the probability integral of the ratio of the minimum of k-1 independent chi-squares to another independent chi-square, all with ν degrees of freedom. The associated distribution theory for this statistic and the tables needed to carry out the procedure R are given in a companion paper [6]. Formulas are obtained for the expected number of populations retained in the selected subset and it is shown that this function attains its maximum when the population variances are all equal. Two generalizations are considered; one deals with the case of unequal degrees of freedom and the other is concerned with a procedure for the selection of the t ($1 \leq t < k$) best populations.

---

*Now at the University of Minnesota

2. Introduction

As alternatives to the classical tests of homogeneity in the Analysis of Variance, new techniques have been developed in recent years which try to incorporate in the original statistical formulation of the problem the plans of the experimenter for further analysis after the hypothesis of homogeneity is tested. For example, if the experimenter tests for homogeneity and regardless of the outcome, ranks his populations on the basis of further analysis of the same data, it would be more realistic to assume at the outset that the parameters are unequal and formulate the main problem as a ranking problem.

The formulation considered in this and earlier papers [3], [4], [5] is that of selecting a subset of k populations which contains the "best" population where the best population is usually defined as the one with the largest (or smallest) parameter value; some further remarks on the motivation of this formulation are contained in these papers.

In the present paper the parameters of interest are the unknown variances of the k normal populations with all means known or all means unknown. The object is to select a subset which includes the population with the smallest variance with a preassigned probability P*, regardless of the true values of the k variances. The procedure R (defined in Section 3) depends only on the sample variances each of which (properly normalized) has a chi-square distribution and hence, under the same formulation, the results of this paper can be applied to any set of k chi-square statistics with a common number of degrees of freedom.

It should be pointed out that this "selecting a subset" formulation is different from and in a certain respect related to the "indifference zone" formulation for the problem of ranking variances treated in [2]. In the latter formulation, an indifference zone in the parameter space is preassigned, the common number of observations needed is tabulated, and the final decision is the selection of a single population which is asserted to be the best population. In the formulation of this paper the number of observations is given, constants needed for carrying out the procedure are tabulated, and the final decision is the selection of a subset of populations which is asserted to contain the best population.

There are several aspects of this formulation which have already been described in the paper dealing with k binomial populations [5] which also apply to this problem. One is that R can be regarded as an elimination or screening procedure. Another is that a confidence statement can be made after experimentation. A third is that the expected size of the selected subset can be regarded as a measure of the efficiency of the procedure.

The main problem is formally described in Section 3 and the procedure R is defined in Section 4. In Section 5 we derive exact and asymptotic expressions for the probability of a correct selection using procedure R and the infimum of this probability over all points in the parameter space. Section 6 deals with the expected number of populations retained in the selected subset. Two generalizations are considered in Section 7, one dealing with the case of unequal degrees of freedom and the other describing a procedure for the problem of selecting a subset containing the t best populations.

3. Formal Statement of the Problem

Let $\Pi_1, \Pi_2, \ldots, \Pi_k$ denote k given normal populations with unknown variances $\sigma_1^2, \sigma_2^2, \ldots, \sigma_k^2$, respectively, (where each $\sigma_i^2 > 0$) and with all means known or all means unknown. The ordered variances are denoted by

$$\sigma^2_{[1]} \leq \sigma^2_{[2]} \leq \ldots. \leq \sigma^2_{[k]}, \tag{3.1}$$

(equalities being allowed for mathematical convenience only). It is assumed that there is no a priori information available about the correct pairing of the k given populations and the ordered scale parameters $\sigma^2_{[i]}$.

The population with variance equal to $\sigma^2_{[1]}$ is called the best population. The goal is to select a subset of the k populations containing the best population. Any such selection will be called a correct selection (CS). Then the problem is to find a rule R such that for a preassigned probability P*

$$P\left\{CS \mid R\right\} \geq P^* , \tag{3.2}$$

regardless of the true unknown values of the populations variances. It is assumed that the same number n of observations will be taken from each population; the case of unequal numbers of observations is briefly considered in Section 7.

From each population Π_i (i = 1, 2, ..., k) we take n observations and, if the mean μ_i is known, we compute the statistic

$$s_i^2 = \frac{1}{n} \sum_{j=1}^{n} (x_{ij}-\mu_i)^2 \qquad (i = 1, 2, \ldots, k) \tag{3.2}$$

which, when multiplied by ν/σ^2 (here $\nu = n$) has the χ^2_ν distribution with ν degrees of freedom. [If the population means are unknown then we use

$$s_i^2 = \frac{1}{n-1} \sum_{j=1}^{n} (x_{ij}-\bar{x})^2 \tag{3.3}$$

and the value of ν is then n-1]. These statistics s_i^2 (i=1, 2, ..., k) form a set of sufficient statistics for the problem and the rule R (defined in Section 4) depends only on these statistics.

It is clear that we prefer rules which make the size S of the selected subset never empty and as small as possible, subject to satisfying (3.2). [One can always attain any specified P^*, even unity, by putting all the populations in the selected subset.]

The rule R is defined in Section 4 for each positive integer $n \geq 2$; let it be denoted here by R(n). Let $\vec{\sigma^2}$ denote the vector of true ordered variances which can vary in a space Ω, in which all vectors have only positive ordered components. Let $\Omega(\delta)$ be that part of Ω in which $\sigma^2_{[1]} \leq \delta\, \sigma^2_{[2]}$ where $\delta > 0$ is preassigned. Finally, let $\epsilon > 0$ be preassigned. Two secondary problems can now be stated. One is to find the smallest common sample size n such that for some particular point $\vec{\sigma^2_o}$ in Ω

$$E\left\{S;\ k,\ \vec{\sigma^2_o},\ P^*,\ R(n)\right\} \leq 1+\epsilon. \tag{3.4}$$

Another problem is to find the smallest common sample size n such that

$$\sup_{\vec{\sigma^2} \text{ in } \Omega(\delta)} E\left\{S;\ k,\ \vec{\sigma^2},\ P^*,\ R(n)\right\} \leq 1+\epsilon. \tag{3.5}$$

The expected size S of the retained subset is regarded as analogous with the complement of the "power" of the test of a hypothesis and both (3.4) and (3.5) are conditions which insure good "power". It is assumed here that both ϵ and δ (or both ϵ and $\sigma_o^{\vec{2}}$) can be specified by the experimenter.

4. Procedure R

Let the ordered values of the k observed sample variances s_i^2 (i = 1, 2, ..., k) all based on a common number ν of degrees of freedom be noted by

$$s_{[1]}^2 \leq s_{[2]}^2 \leq \ldots \leq s_{[k]}^2. \tag{4.1}$$

The procedure R is then defined as follows. Procedure R: "Retain Π_i in the selected subset if and only if

$$s_i^2 \leq s_{[1]}^2/c \tag{4.2}$$

where $c = c(\nu, k, P^*)$ is a constant with $0 < c \leq 1$ which is determined in advance of experimentation."

The constant c is chosen to be the largest value which satisfies the basic probability requirement (3.2) for all true configurations $\sigma^{\vec{2}} = (\sigma_1^2, \sigma_2^2, \ldots, \sigma_k^2)$. Tables of c-values for $\nu = 2(2)50$, $k = 2(1)11$ and $P^* = \cdot 75, \cdot 90, \cdot 95$, and $\cdot 99$ are given in a compansion paper [6]; these c-values can also be regarded as percentage points of a smallest Studentized χ^2-statistic as explained in Section 5.

Illustration

From each of k=5 populations Π_i, a statistic T_i is computed such that T_i/θ_i has a χ_ν^2 distribution where the common value of ν is 14. It is desired

to have a $P\{CS \mid R\}$ of at least $P^* = \cdot 95$ for any true point in the parameter space. From Table IB of [6] the appropriate c-value is $c(14, 5, \cdot 95) = .2911$. The observed values of T_i are 7.12, 2.02, 3.09, 5.05, 7.28 so that $T_{[1]} = 2.02$ and $T_{[1]}/c = 6.94$. Applying the procedure R, we find that the populations in the selected subset are those three that gave rise to the values 2.02, 3.09, and 5.05. At this point the experimenter can assert with confidence level $\cdot 95$ that one of these three populations has the smallest value of θ. In the case of ranking variances of normal populations $\theta_i = \sigma_i^2$ is the population variance of Π_i and $T_i = \nu s_i^2$ is proportional to the sample variance from Π_i $(i = 1, 2, \ldots, 5)$.

5. The Probability of a Correct Selection

In this section exact and asymptotic expressions are derived for the probability of a correct selection for general k, ν, and any point in the parameter space and also for the infimum of this probability over all points in the parameter space. It will be convenient here to associate ν_i degrees of freedom (the ν_i need not be equal) with the sample variance s_i^2 from Π_i $(i = 1, 2, \ldots, k)$ and to allow the ν_i to take on even as well as odd integer values; this generalization is discussed in Section 7. Under this general framework Section 5.2 points out that the $P\{CS \mid R\}$ approaches its infimum as the variances approach equality (with $\sigma_{[1]}^2 = \sigma_{[2]}^2$ only in the limit). It should be noted that the infimum is not the same as the value of the $P\{CS \mid R\}$ when the variances are equal (the latter has not been defined but by any reasonable extension of our definition should be unity); mathematically, we can use the configuration with all variances equal provided we "tag" a particular one of the populations and regard it as being the best. In Section 5.3 the $P\{CS \mid R\}$ is shown to be equivalent to the complement of the c.d.f. of a studentized smallest chi-square statistic and this leads to an approximation of the $P\{CS \mid R\}$ based on asymptotic normality.

5.1 Exact Expressions for the Probability of a Correct Selection

Let $s^2_{(i)}$ denote the (unknown) sample variance that is associated with the i^{th} smallest population variance, $\sigma^2_{[i]}$; let $\nu_{(i)}$ denote the number of degrees of freedom associated with $s^2_{(i)}$. The procedure R described in Section 3 yields a correct selection if and only if the event

$$s^2_{(1)} \leq \frac{1}{c} \operatorname{Min}_{\alpha} s^2_{(\alpha)} \qquad (\alpha = 1, 2, \ldots, k) \tag{5.1}$$

occurs. Since $0 < c \leq 1$, the occurence of the event (5.1) is equivalent to the occurence of the event

$$s^2_{(1)} \leq \frac{1}{c} \operatorname{Min}_{\alpha \neq 1} s^2_{(\alpha)} \qquad (\alpha = 2, 3, \ldots, k). \tag{5.2}$$

Hence, the probability of a correct selection is given by

$$P\left\{CS \mid R\right\} = P\left\{s^2_{(1)} \leq \frac{1}{c} \operatorname{Min}_{\alpha \neq 1} s^2_{(\alpha)} \ (\alpha = 2, 3, \ldots, k)\right\} \tag{5.3}$$

$$= P\left\{\frac{\nu_{(\alpha)}\, s^2_{(\alpha)}}{\sigma^2_{[\alpha]}} \geq \frac{c\ \nu_{(\alpha)}\ \sigma^2_{[1]}}{\nu_{(1)}\ \sigma^2_{[\alpha]}} \left(\frac{\nu_{(1)}\, s^2_{(1)}}{\sigma^2_{[1]}}\right) (\alpha = 2, 3, \ldots, k)\right\}$$

$$= \int_0^\infty g_\nu(x) \prod_{\alpha=2}^{k} \left[1 - G_\nu\left(\frac{c\ \nu_{(\alpha)}\ \sigma^2_{[1]}}{\nu_{(1)}\ \sigma^2_{[\alpha]}}\, x\right)\right] dx$$

where $G_\nu(x)$ and $g_\nu(x)$ are the chi-square c·d·f· and p·d·f· with ν degrees of freedom, respectively; we can also regard $G_\nu(x)$ and $g_\nu(x)$ as the gamma c·d·f· and p·d·f·, respectively.

Suppose in the next to the last expression of (5.3) we make the transformation

$$(5.4) \qquad u_1 = \frac{\nu_{(1)}\, s^2_{(1)}}{\sigma^2_{[1]}}\,, \quad u_\alpha = \frac{\nu_{(\alpha)}\, \sigma^2_{[1]}\, s^2_{(\alpha)}}{\nu_{(1)}\, \sigma^2_{(\alpha)}\, s^2_{(1)}} \quad (\alpha = 2,\ 3,\ \ldots,\ k)$$

then the limits of u_1 are from 0 to ∞ and upon integrating out u_1, we obtain the (k-1)-fold integral

$$(5.5) \qquad P\left\{CS \mid R\right\} = \int_{\frac{cm_{(2)}\,\sigma^2_{[1]}}{m_{(1)}\,\sigma^2_{[2]}}}^{\infty} \cdots \int_{\frac{cm_{(k)}\,\sigma^2_{[1]}}{m_{(1)}\,\sigma^2_{[k]}}}^{\infty} \frac{\Gamma(m_1+\ldots+m_k)}{\prod_{\alpha=1}^{k}\Gamma(m_\alpha)} \frac{\prod_{\alpha=2}^{k}\left[u_\alpha^{m_\alpha-1}\, du_\alpha\right]}{(1+u_2+\ldots+u_k)^{m_1+\ldots+m_k}}$$

where $m_{(\alpha)} = \nu_{(\alpha)}/2$.

5.2 Exact Expressions for the Infimum of the $P\left\{CS \mid R\right\}$

It follows from both (5.3) and (5.5) that, for fixed c and fixed $m_{(\alpha)}$ (or $\nu_{(\alpha)}$), the $P\left\{CS \mid R\right\}$ depends only on the ratios of the variances and that it approaches its infimum by setting $\sigma^2_{[k]} = \sigma^2_{[k-1]} = \ldots = \sigma^2_{[2]}$ letting $\sigma^2_{[2]} \to \sigma^2_{[1]}$, with equality only in the limit. Hence, we obtain for the case of a common m [letting $\inf_\Omega$ denote the infimum over all points in the parameter space $\sigma_i^2 > 0$ $(i = 1, 2, \ldots, k)$],

$$(5.6) \qquad \inf_\Omega P\left\{CS \mid R\right\} = \int_c^\infty \cdots \int_c^\infty \frac{\Gamma(km)}{[\Gamma(m)]^k} \frac{\prod_{\alpha=2}^{k}[u_\alpha^{m-1}\, du_\alpha]}{(1+u_2+\ldots+u_k)^{km}}$$

$$= \int_0^\infty \left[1-G_\nu(cx)\right]^{k-1} g_\nu(x)\, dx\,.$$

In the special case when ν is an even integer (so that m is an integer) further simplification takes place in the last expression of (5.6) and we can obtain a result in the form of a finite series. This result is described in the companion paper [6], where it is used to construct tables for the largest c-value satisfying the probability requirement (3.2).

5.3 Approximation to Inf $P\left\{CS \mid R\right\}$ Based on Asymptotic Normality

Let $\chi_j^2(\nu)$ ($j = 0, 1, 2, \ldots, p$; $p = k-1$) denote k independent chi-square chance variables with a common number of degrees of freedom ν. It follows from (5.2) and the fact that the $P\left\{CS \mid R\right\}$ approaches its infimum as the variances approach equality that we can write the basic probability requirement in the form

$$P\left\{\frac{\underset{j=1, \ldots, p}{\text{Min}} \chi_j^2(\nu)}{\chi_o^2(\nu)} \geq c\right\} = P^* \quad . \tag{5.7}$$

Hence, the determination of c to satisfy (3.2) for all points in the parameter space is equivalent to the determination of a lower percentage points of the Studentized smallest chi-square statistic with ν degrees of freedom for all k chi-squares.

Using the fact that for large ν the statistic $Y_j = \log (\chi_j^2(\nu)/\nu)$ tends to normality as $\nu \to \infty$ and also the fact that the set of statistics $\left\{Y_j - Y_o\right\}$ tends to a joint p-variate normal distribution, we obtain (the details are similar to those used in Section 5 of [2] for the case t=1) the approximation

$$\inf_{\Omega} P\left\{CS \mid R\right\} \cong \int_{-\infty}^{\infty} \left[1-F(x-d)\right]^{k-1} f(x)\, dx \tag{5.8}$$

$$= \int_{-\infty}^{\infty} \left[F(x+d)\right]^{k-1} f(x)\, dx$$

where F(x) and f(x) are the standard Normal c·d·f·and p·d·f·, respectively, and

$$d = \sqrt{(\nu-1)/2}\, \log (1/c). \tag{5.9}$$

Another expression for (5.8) is

$$\inf_{\Omega} P\left\{CS \mid R\right\} = (k-1)\int_{-\infty}^{\infty} F(x+d)\left[1-F(x)\right]^{k-2} f(x)\,dx, \tag{5.10}$$

and this is the form used in [7] where this quantity is extensively tabulated.

For large values of ν an approximate solution to the largest value of c satisfying the probability requirement (3.2) can be obtained by equating the last expression in (5.8) to P^* and using (5.9) to solve for c. Values in d are tabulated in Table I of [1] for k = 2(1)10 and many values of P^* [our d corresponds to $\lambda\sqrt{N}$ in his notation] and also in Table AI of [3] for k = 2(1)51 for selected values of P^* [our p and P^* correspond to his n and $1-\alpha$, respectively].

To illustrate numerically the closeness of the normal approximation we take a c-value out of Table IC of [6] corresponding to $P^* = \cdot 95$, k = 3, $\nu = 50$, namely, c = .5761, which is based on an exact calculation, setting the last member of (5.6) equal to $P^* = \cdot 95$ and solving for c This is to be compared with a computation based on (5.8) for which we first must compute d, obtaining

$$d = \sqrt{\frac{\nu-1}{2}}\,\log_e(1/c) = \frac{7\sqrt{2}}{2}\left(\frac{1}{.5761}\right) = 2.729. \tag{5.11}$$

From the table headed P(1,3) in [7], we obtain by linear interpolation. P = .9515, which is slightly greater than the correct value, namely $P^* = \cdot 95$.

6. Expected Size of the Selected Subset

For the procedure R the size S of the selected subset is a chance variable which can take on only integer values from 1 to k, inclusive. For any fixed values of ν, k, and P^*, the expected size of the selected subset is a function of the true configuration $\vec{\sigma^2} = \left\{\sigma_1^2, \sigma_2^2, \ldots, \sigma_k^2\right\}$ and this function can be regarded as a criterion of the efficiency of any procedure which satisfies the basic probability requirement (3.2). In analogy with power function considerations, one secondary problem is to find the smallest common sample size n necessary to control $E\left\{S\right\}$ at some preassigned level for a particular alternative in the parameter space; alternatively, we may wish to control the maximum of $E\left\{S\right\}$ over all parameter points in the subset $\Omega(\delta)$ of Ω given by $\sigma^2_{[i]} \geq \delta\, \sigma^2_{[1]}$ $(i = 2, 3, \ldots, k)$ with $\delta > 1$.

6.1 Exact Expression for the Expected Size

Let Y_i denote a chance variable which equals 1 if Π_i is included in the selected subset and equals 0 otherwise. Then $S = \sum_{i=1}^{k} Y_i$ and hence

$$(6.1) \qquad E\left\{S\right\} = E\left\{\sum_{i=1}^{k} Y_i\right\} = \sum_{i=1}^{k} E\left\{Y_i\right\}$$

$$= \sum_{i=1}^{k} P\left\{\Pi_i \text{ is included in the selected subset}\right\}$$

for any values of ν, k, P^*, and $\vec{\sigma^2}$ [these symbols are suppressed in (6.1)].

Using an argument similar to the one in Section 5.1 for obtaining the exact $P\left\{CS\right\}$ we obtain from (6.1)

$$(6.2) \qquad E\left\{S\right\} = \sum_{i=1}^{k} \int_0^{\infty} g_\nu(x) \prod_{\substack{j=1 \\ j\neq i}}^{k} \left[1 - G_\nu\left(\theta'_{ij} c x\right)\right] dx$$

where $\theta'_{ij} = \sigma_i^2/\sigma_j^2$. [It should be noted that if $\theta_{ij} = \sigma^2_{[i]}/\sigma^2_{[j]}$ is used in (6.2), then the i^{th} term on the right hand side of (6.2) is the probability that the population associated with $\sigma^2_{[i]}$ is included in the selected subset and hence the sum is again $E\{S\}$.]

The particular configuration

$$\delta\,\sigma^2_{[1]} = \sigma^2_{[\alpha]},\ (\alpha = 2, 3, \ldots, k) \tag{6.3}$$

is of some special interest and in this case (6.2) reduces to

$$E\{S\} = \int_0^\infty \left[1-G_\nu\left(\frac{cx}{\delta}\right)\right]^{k-1} g_\nu(x)\,dx + (k-1)\int_0^\infty \left[1-G_\nu(c\delta x)\right]\left[1-G_\nu(cx)\right]^{k-2} g_\nu(x)\,dx \tag{6.4}$$

where the first term on the right side of (6.4) is the $P\{CS\}$ for the configuration (6.3). In particular if all the variances are equal then $\delta = 1$ and $E\{S\}$ equals $k\,P^*$.

6.2 Maximum Value of $E\{S\}$

It will now be shown that the maximum value of $E\{S\}$ takes place when all the population variances are equal. If we set the m largest variances ($1 \leq m < k$) equal to a common value σ^2 (say) and define $\theta_{i.} = \sigma^2_{[i]}/\sigma^2 = \theta_{.i}^{-1}$ then writing $Q = E\{S \mid \sigma^2_{[k]} = \sigma^2_{[k-1]} = \ldots = \sigma^2_{[k-m+1]} = \sigma^2\}$, we obtain from the remark following 6.2,

$$(6.5)\qquad Q = \sum_{i=1}^{k-m} \int_0^\infty \left[1-G_\nu(\theta_{i\cdot}cx)\right]^m g_\nu(x) \prod_{\substack{j=1\\ j\neq i}}^{k-m} \left[1-G_\nu(\theta_{ij}cx)\right] dx$$

$$+ m \int_0^\infty \left[1-G_\nu(cx)\right]^{m-1} g_\nu(x) \prod_{j=1}^{k-m} \left[1-G_\nu(\theta_{\cdot j}cx)\right] dx \ .$$

We now show that the right hand member of (6.5) is a non-increasing function of σ^2 for $\frac{1}{k} \leq P^* \leq 1$ (actually the proof shows that it is a strictly decreasing function of σ^2 for $\frac{1}{k} < P^* < 1$). This proves that it is a maximum when $\sigma^2 = \sigma^2_{[k-m]}$ and since this holds for any integer $m < k$, the desired result will follow.

To show that Q is monotonic we differentiate Q with respect to σ^2 and show that the result is negative for $\frac{1}{k} < P^* < 1$ and $\sigma^2 > \sigma^2_{[k-m]}$. Differentiation gives

$$(6.6)\qquad \frac{dQ}{d\sigma^2} = mc \sum_{i=1}^{k-m} \left\{ \frac{\sigma^2_{[i]}}{\sigma^4} \int_0^\infty x \left[1-G_\nu(\theta_{i\cdot}cx)\right]^{m-1} g_\nu(\theta_{i\cdot}cx)\, g_\nu(x) \prod_{\substack{j=1\\ j\neq i}}^{k-m} \left[1-G_\nu(\theta_{ij}cx)\right] dx \right.$$

$$\left. - \frac{1}{\sigma^2_{[i]}} \int_0^\infty x \left[1-G_\nu(cx)\right]^{m-1} g_\nu(\theta_{\cdot i}cx)\, g_\nu(x) \prod_{\substack{j=1\\ j\neq i}}^{k-m} \left[1-G_\nu(\theta_{\cdot j}cx)\right] dx \right\} .$$

For $P^* < 1$ it is clear that $c > 0$. If we let $x = x'\theta_{i\cdot}$ in the second integral and drop primes then (6.6) becomes

$$(6.7)\qquad \frac{dQ}{d\sigma^2} = mc \sum_{i=1}^{k-m} \frac{\sigma^2_{[i]}}{\sigma^2} \int_0^\infty x \left[1-G_\nu(\theta_{i\cdot}cx)\right]^{m-1} \prod_{\substack{j=1\\ j\neq i}}^{k-m} \left[1-G_\nu(\theta_{ij}cx)\right] \cdot$$

$$\left\{ g_\nu(\theta_{i\cdot}cx)\, g_\nu(x) - g_\nu(cx)\, g_\nu(\theta_{i\cdot}x) \right\} dx \ .$$

It suffices to prove that the expression in braces in (6.7) is negative for each $x > 0$. Treating $g_\nu(x)$ as a gamma density, we obtain

$$\log\left\{g_\nu(\theta_{i.}cx)\; g_\nu(x) \div g_\nu(cx)\; g_\nu(\theta_{i.}x)\right\} = -(1-c)\;(1-\theta_{i.}).x. \tag{6.8}$$

Since $c < 1$ for $P^* > \frac{1}{k}$ it follows from (6.8) that for $\sigma^2 > \sigma^2_{[k-m]}$ the expected size of the selected subset is a decreasing function of σ^2. This proves the desired result.

It follows that

$$\max_{\Omega} E\left\{S\right\} = k\int_0^\infty [1-G_\nu(cx)]^{k-1}\; g_\nu(x)\; dx = k\; P^* \tag{6.9}$$

which does not depend on the common value of σ^2.

The above proof also shows that in the subset $\Omega(\delta)$ of the parameter space the function $E\left\{S\right\}$ takes on its maximum value when (6.3) holds. Hence, the maximum value is given by (6.4). If it is desired to find the smallest value of the common sample size n (equal to ν or ν-1) such that $E\left\{S\right\} \leq 1+\epsilon$ for all points in $\Omega(\delta)$ ($\epsilon > 0$ and $\delta > 1$ are preassigned) then the solution in ν can be obtained by equating the right hand member of (6.4) to $1+\epsilon$.

6.3 An Approximation to $E\left\{S\right\}$ Based on Asymptotic Normality

Using the fact that $\log\left(s^2_{(i)}/\sigma^2_{[i]}\right)$ $(i = 1, 2, \ldots, k)$ are independently and asympototically normally distributed with a common mean and a common variance $2/(\nu-1)$, we now derive the following approximation for $E\left\{S\right\}$,

$$E\left\{S\right\} = \sum_{i=1}^{k}\int_{-\infty}^{\infty}\prod_{\substack{j=1\\ j\neq i}}^{k}\left[1-F\left(x+\frac{\log(c\theta_{ij})}{\sqrt{\frac{2}{\nu-1}}}\right)\right] f(x)\; dx. \tag{6.10}$$

For particular configuration (6.3) this reduces to

$$(6.11) \qquad E\{S\} = \int_{-\infty}^{\infty} \left[1-F\ (x-d-\delta_1)\right]^{k-1} f(x)\ dx$$

$$+ (k-1) \int_{-\infty}^{\infty} \left[1-F\ (x-d)\right]^{k-2} \left[1-F\ (x-d+\delta_1)\right] f(x)\ dx\ ,$$

where d is given by (5.9) and $\delta_1 = \sqrt{\frac{\nu-1}{2}}\ \log \delta$.

7. Generalizations

In this section we consider two different directions for generalizing the problem treated above. One generalization is to allow unequal degrees of freedom associated with each of the sample variances from the k given populations. The other generalization deals with the problem of selecting a subset containing the t "best" populations, i.e., the populations with the t smallest variances, for $t > 1$. For each of these two generalizations, a general conjecture will be made about the worst configuration and each conjecture will be proved for at least one interesting special case. The appropriate value of c to meet a specified probability requirement has not been computed for either of these generalizations.

7.1 Unequal Degrees of Freedom

In order to find the infimum of the probability of a correct selection, we have to consider both changing the values of the unknown population variances and also different possible pairings of the known degrees of freedom $\nu_1, \nu_2, \ldots, \nu_k$ with the ordered population variances. The first part, i.e., changing the values of the population variances, was taken into account in Section 5.2 and it follows from there that we need only consider the case of equal variances (or, more precisely, a sequence of configurations with the smallest population variance

different from the others and with all population variances approaching equality). For the second part, it is clear from (5.3) that we are only concerned with at most k of k! possible pairings since (5.3) depends only on which ν is associated with the smallest population variance. We now state a general conjecture about these k possible associations, using the notation of Section 5.2.

Conjecture 1: For any fixed set of k values of ν (say, $\nu_1, \nu_2, \ldots, \nu_k$) the probability of a correct selection $P\left\{CS \mid R\right\}$ is decreased by an interchange of $\nu_{(1)}$ and $\nu_{(\alpha)}$ $(\alpha > 1)$ if $\nu_{(1)} < \nu_{(\alpha)}$.

[This conjecture can be regarded as the result of two separate conjectures; the first is that the $P\left\{CS \mid R\right\}$ is decreased by increasing $\nu_{(1)}$ holding all the remaining ν-values fixed and the second is that for each $\alpha > 1$ the $P\left\{CS \mid R\right\}$ is decreased by decreasing $\nu_{(\alpha)}$ holding all the remaining ν-values fixed.]

A proof of this conjecture is given at this point for the following special cases:

Case 1: All k statistics s_i^2 have 2 degrees of freedom, except for one of them which has ν degrees of freedom, where $\nu \geq 2$ is an even integer; let $m = \nu/2$.

Case 2: There are two populations (k = 2) and ν_1, ν_2 are any integers.

Proof for Case 1:

If we associate m with the best population then using (5.6) the probability of a correct selection becomes

$$(7.1) \qquad \int_0^\infty \frac{x^{m-1}e^{-x}}{\Gamma(m)} \left[e^{-\frac{cx}{m}}\right]^{k-1} dx = \left[1 + \frac{c(k-1)}{m}\right]^{-m}$$

and if we associate m with any of the other populations then, again using (5.6), the probability of a correct selection becomes

$$(7.2)\qquad \int_0^\infty [e^{-cx}]^{k-2}\, e^{-x}\left[e^{-cmx}\sum_{j=0}^{m-1}\frac{(cmx)^j}{j!}\right]dx = \sum_{j=0}^{m-1}\frac{(cm)^j}{\left[1+c(k+m-2)\right]^{j+1}}$$

$$= \frac{1-\left[\frac{cm}{1+c(k+m-2)}\right]^m}{1+c(k-2)}$$

Hence, to prove that the conjecture holds in this case it must be shown that for any c ($0 \leq c \leq 1$) and any integer k ($k \geq 2$) the last member of (7.1) is not greater than the last member of (7.2). Since equality holds for $c = 0$, it is sufficient to show that for $c > 0$ and any integer $k \geq 2$

$$(7.3)\qquad \left[1+c(k-2)\right]\left[1+\frac{c(k-1)}{m}\right]^{-m} + \left[1+\frac{1+c(k-2)}{cm}\right]^{-m} \leq 1.$$

It is well known (and easy to show) that the function $\left(1+\frac{\Delta}{x}\right)^{-x}$, for any fixed $\Delta > 0$, is a strictly decreasing function of x. It follows that each of the terms on the left side of (7.3) must take its maximum value at $m = 1$. Since equality clearly holds in (7.3) for $m = 1$, the result is proved.

Proof for Case 2:

Assume that $\nu_1 \leq \nu_2$. It follows from (5.2) that proving the desired result is equivalent to showing that for any $c (0 < c \leq 1)$

$$(7.4)\qquad P\left\{F_{\nu_1,\,\nu_2} \leq \frac{1}{c}\right\} \geq P\left\{F_{\nu_2,\,\nu_1} \leq \frac{1}{c}\right\}$$

or equivalently that

$$(7.5) \qquad P\left\{F_{\nu_1,\,\nu_2} > \frac{1}{c}\right\} \leq P\left\{F_{\nu_2,\,\nu_1} > \frac{1}{c}\right\}$$

where $F_{\nu_1,\,\nu_2}$ is the usual (Snedecor) F chance variable with ν_1 and ν_2 d.f. Since $1/c > 1$, it is sufficient to show that the densities corresponding to the left and right members of (7.5) satisfy the same inequality for each $F > 1$, i.e., that for $F > 1$ and $\nu_1 \leq \nu_2$

$$(7.6) \qquad \frac{F^{(\nu_1-2)/2}}{(\nu_2+\nu_1 F)^{(\nu_1+\nu_2)/2}} \leq \frac{F^{(\nu_2-2)/2}}{(\nu_1+\nu_2 F)^{(\nu_1+\nu_2)/2}}$$

By algebraic simplification, this can also be written as

$$(7.7) \qquad \frac{F^{\nu_1'} - F^{-\nu_1'}}{\nu_1'} \leq \frac{F^{\nu_2'} - F^{-\nu_2'}}{\nu_2'}$$

where $\nu_1' = \nu_1/(\nu_1+\nu_2)$ and $\nu_2' = \nu_2/(\nu_1+\nu_2)$. Since $\nu_1' \leq \nu_2'$, it is sufficient to show that the function $Q(x) = (F^x - F^{-x})/x$ with $F > 1$ is a strictly increasing function of x for $0 < x < 1$. Letting $Q'(x)$ denote $\frac{dQ(x)}{dx}$ and $Q_1(x) = x^2 Q'(x)$, we obtain by differentiation

$$(7.8) \qquad Q_1(x) = x\,(F^x + F^{-x}) \log F - (F^x - F^{-x})$$

$$(7.9) \qquad Q_1'(x) = x\,(F^x - F^{-x})\,(\log F)^2 > 0,$$

using the fact that $x > 0$ and $F > 1$ in the last inequality. Since $Q_1(0) = 0$, it follows from (7.9) that $Q_1(x) > 0$ for $0 < x < 1$ and the same result then holds for $Q'(x)$. It follows that $Q(x)$ is strictly increasing for $0 < x < 1$, as was to be shown.

7.2 Selecting a Subset Containing the t Best Populations

The second generalization deals with the problem of selecting a subset of the k given populations containing the populations with the t smallest variances (with fixed $t \leq k$). In this generalization we will restrict our attention to the case of equal degrees of freedom, for the sake of simplicity. The proposed procedure for this problem is to order the sample variances, (let $s^2_{[i]}$ denote the i^{th} smallest and let s_i^2 denote the one from population Π_i as in Section 3), and put Π_i in the selected subset if and only if

$$s_i^2 \leq \frac{s^2_{[t]}}{c} \tag{7.10}$$

where the constant c with $0 < c < 1$ is evaluated in the "worst configuration" to satisfy a prescribed probability requirement. Then the selected subset must contain at least t populations and if the sample size is too small and/or the population variances are too "close together" then the selected subset may even contain all k populations.

To solve this problem one mustfirst ascertain the worst configuration, i.e., find a sequence of vectors (whose components are population variances) for which the probability of a correct selection approaches its infimum. The above is not easy, principally because of the difficulty in obtaining a simple expression for the probability of a correct selection. We now state a general conjecture about the worst configuration.

Conjecture 2: For any fixed t ($1 \leq t < k$), the $P\left\{CS \mid R_t\right\}$ approaches its infimum if the k-s+1 largest population variances ($1 \leq s \leq t$) approach equality (with equality of $\sigma^2_{[t]}$ and $\sigma^2_{[t+1]}$ only in the limit) and the remaining s-1 population variances approach (or are set equal to) zero. The integer s takes all values between 1 and t and is a non-decreasing function of the specified probability P^*; in particular, for P^* close to unity we set t-1 population variances equal to zero and the tables for the case t = 1 (with the k-value taken to be k-t+1) can then be used to determine the appropriate c-value.

It is interesting to note that for t = 1 we must have s = 1 and the conjecture is consistent with the results of Section 5.2. A proof of the conjecture will be given for the special case k = 3, t = 2, and a common $\nu = 2$. In order to prove that the conjecture holds for this special case, it is necessary to obtain an expression for the probability of a correct selection, which we now derive for general t.

We define the integer chance variable T′ ($1 \leq T' \leq k$) by saying that $T' = t'$ whenever $s^2_{(t')} \equiv s^2_{[t]}$, i.e., when $s^2_{[t]}$ comes from the population with variance $\sigma^2_{[t']}$. We define the integer chance variable R ($0 \leq R \leq k-t$) for fixed c by saying that R = r whenever

$$s^2_{[t+r]} < \frac{s^2_{[t]}}{c} < s^2_{[t+r+1]} \tag{7.11}$$

where $s^2_{[k+1]} = \infty$ and we disregard all possibilities of equality. Then $s^2_{[t]}$ and $s^2_{[t]}/c$ divide the k-1 remaining sample variances into 3 sets S_1, S_2, S_3 of sizes t-1, r, and k-t-r, respectively. Let $P_1\left\{(S_{1\alpha}, S_{2\alpha}, S_{3\alpha}), t', r; c\right\}$ denote the probability of a correct selection when $T' = t' \leq t$, R = r and $(S_{1\alpha}, S_{2\alpha}, S_{3\alpha})$

represents a partition of the set of t-1 "smaller" sample variances and k-t "larger" sample variances that corresponds to a correct decision; let $P\left\{(S_{1\beta}, S_{2\beta}, S_{3\beta}) , t', r; c\right\}$ denote the same for $T' = t' > t$. Then the probability of a correct selection (letting R_t denote the proposed procedure) can be written as

$$P\left\{CS \mid R_t\right\} = \sum_{t'=1}^{t} \sum_{r=0}^{k-t} \frac{\binom{k-t}{r}\binom{t'+r-1}{t-1}}{\sum_{\alpha=1}} P_1\left\{(S_{1\alpha}, S_{2\alpha}, S_{3\alpha}) , t', r; c\right\}$$

(7.12)

$$+ \sum_{t'=t+1}^{k} \sum_{r=1}^{k-t} \frac{\binom{k-t-1}{r-1}\binom{t'+r-1}{t-1}}{\sum_{\beta=1}} P_2\left\{(S_{1\beta}, S_{2\beta}, S_{3\beta}) , t', r; c\right\} .$$

If $S_{1\alpha}$ consists of the sample variances $s^2_{(i)}$ for $i = (i_1, i_2, \ldots, i_{t-1})$, $S_{2\alpha}$ consists of the $s^2_{(i)}$ for $i = (i_{t+1}, i_{t+2}, \ldots, i_{t+r})$ and $S_{3\alpha}$ consists of the $s^2_{(i)}$ for $i = (i_{t+r+1}, i_{t+r+2}, \ldots, i_k)$ and if we let $i_t = t'$ then we can write

$$P_1\left\{(S_{1\alpha}, S_{2\alpha}, S_{3\alpha}), t', r; c\right\} = \int_0^\infty \prod_{j=1}^{t-1} \left[G_\nu\left(\frac{\sigma^2_{[t']}}{\sigma^2_{[i_j]}} x\right)\right] .$$

(7.13)

$$\cdot \prod_{j=t+1}^{t+r} \left[G_\nu\left(\frac{\sigma^2_{[t']}}{\sigma^2_{[i_j]}} \frac{x}{c}\right) - G_\nu\left(\frac{\sigma^2_{[t']}}{\sigma^2_{[i_j]}} x\right)\right] \prod_{j=t+r+1}^{k} \left[1 - G_\nu\left(\frac{\sigma^2_{[t']}}{\sigma^2_{[i_j]}} x\right)\right] g_\nu(x)\, dx$$

where $G_\nu(x)$ and $g_\nu(x)$ are the chi-square (χ^2) c·d·f· and density, respectively, for ν degrees of freedom. The same expression (7.13) also holds for

$P_2\left\{(S_{1\beta}, S_{2\beta}, S_{3\beta}), t', r; c\right\}$ except that $t' > t$ and hence we now consider partitions of the set of t "small" variances and k-t-1 "large" variances that correspond to a correct selection. Some simplification in the first of the two expressions of (7.12) can be obtained by considering separately those cases in which the t best populations yield the t smallest sample variances. The first expression of (7.12) then breaks up into the two parts

$$\sum_{t'=1}^{t} \int_0^\infty \prod_{\substack{i=1 \\ i\neq t'}}^{t} \left[G_\nu\left(\frac{\sigma^2_{[t']}}{\sigma^2_{[i]}} x\right)\right] \prod_{i=t+1}^{k} \left[1-G_\nu\left(\frac{\sigma^2_{[t']}}{\sigma^2_{[i]}} x\right)\right] g_\nu(x)\, dx$$

(7.14)

$$+ \sum_{t'=1}^{t} \sum_{r=1}^{k-t} \sum_{\alpha=1}^{\binom{k-t}{r}\left[\binom{t+r-1}{t-1} - 1\right]} P_1\left\{(S_{1\alpha}, S_{2\alpha}, S_{3\alpha}), t', r; c\right\}$$

where the first part does not depend on c or r and in the second part the index α sums over all the remaining partitions in which at least one ($r > 0$) of k-t populations with larger population variances is associated with one of the t-1 smallest observed sample variances.

It should also be noted that if the conjecture is correct then we can set all σ^2's equal in (7.13) and in (7.14) to obtain a certain amout of further simplification.

Proof of Conjecture for Special Case k=3, t=2, and Common ν=2

Suppose we consider the special case k=3, t=2, and a common ν. Then by (7.13) and (7.14) we have

$$P\left\{CS \mid R_2\right\} = P\left\{s^2_{(1)} < s^2_{(2)} < s^2_{(3)}\right\} + P\left\{s^2_{(2)} < s^2_{(1)} < s^2_{(3)}\right\}$$

$$(7.15) \quad + P\left\{s^2_{(3)} < s^2_{(1)} < s^2_{(2)} < \tfrac{1}{c}\, s^2_{(1)}\right\} + P\left\{s^2_{(3)} < s^2_{(2)} < s^2_{(1)} < \tfrac{1}{c}\, s^2_{(2)}\right\}$$

$$+ P\left\{s^2_{(1)} < s^2_{(3)} < s^2_{(2)} < \tfrac{1}{c}\, s^2_{(3)}\right\} + P\left\{s^2_{(2)} < s^2_{(3)} < s^2_{(1)} < \tfrac{1}{c}\, s^2_{(3)}\right\}$$

where the 1st, 2nd, and 3rd pair of terms corresponds to the 1st part of (7.14), the 2nd part of (7.14) and the 2nd part of (7.12), respectively. Note that there are symmetries present between the odd and even terms and also between the second and third pair of terms.

Since t is so close to k we gain a little extra simplification by taking complements and writing

$$P\left\{CS \mid R_2\right\} = 1 - \left[P\left\{s^2_{(3)} < s^2_{(1)} < \frac{s^2_{(1)}}{c} < s^2_{(2)}\right\} + P\left\{s^2_{(3)} < s^2_{(2)} < \frac{s^2_{(2)}}{c} < s^2_{(1)}\right\}\right.$$

$$(7.16) \quad \left. - \left[P\left\{s^2_{(1)} < s^2_{(3)} < \frac{s^2_{(3)}}{c} < s^2_{(2)}\right\} + P\left\{s^2_{(2)} < s^2_{(3)} < \frac{s^2_{(3)}}{c} < s^2_{(1)}\right\}\right]\right.$$

For $\nu=2$ this is easily computed and we obtain

$$P\left\{CS \mid R_2\right\} = 1 - \frac{cx}{1+cx} - \frac{c}{x+c} + \frac{c(1+xy)}{c+y+cxy} - \frac{c}{c+y} + \frac{c(1+y)}{c+xy+cy} - \frac{c}{c+xy}$$

$$(7.17) \quad = \frac{x(1-c^2)}{(c+x)\ (1+cx)} + \frac{cxy^2}{(c+y)\ (c+y+cxy)} + \frac{cxy^2}{(c+xy)\ (c+xy+cy)}$$

where $x = \sigma^2_{[2]}/\sigma^2_{[1]}$ and $y = \sigma^2_{[3]}/\sigma^2_{[2]}$ are both equal to or greater than unity and $0 < c \leq 1$. Since the first term above is independent of y and each of the last two are strictly increasing in y it follows that for any x we obtain a minimum by setting $y = 1$, obtaining

$$(7.18)\qquad P(x \mid y=1) = \frac{x(1-c^2)}{(c+x)(1+cx)} + \frac{cx}{(c+1)(c+1+cx)} + \frac{cx}{(c+x)(2c+x)},$$

where the symbol $P(x \mid y=1)$ is used to denote $P\left\{CS \mid R_2\right\}$ for $y=1$.

We now note that at the extreme values of x we have

$$(7.19)\qquad P(1 \mid y=1) = \frac{3c+1-2c^2}{(1+c)(2c+1)}$$

$$(7.20)\qquad P(\infty \mid y=1) = \frac{1}{1+c}$$

and it is easily verified that for $c = \frac{1}{2}$ these two functions are equal with common value $\frac{2}{3}$ and that

$$(7.21)\qquad P(\infty \mid y=1) \lesseqgtr P(1 \mid y=1) \text{ for } c \lesseqgtr \frac{1}{2}.$$

We now show that for $c \geq \frac{1}{2}$ the minimum for all $x \geq 1$ is given by (7.19) and for $c \leq \frac{1}{2}$ the minimum for all $x \geq 1$ is given by (7.20). In the first part it is sufficient to show that for $x \geq 1$ the difference between (7.18) and (7.19) is non-negative. After some tedious algebra, this difference can be written as

$$(7.22)\qquad \frac{c^2(x-1)\left\{\left[2(c+1)^2 + 4c(1-c^2)\right] + 2x(1+4c^2+2c^3) + x^2(2c-1)(3c^2+2c+2) + x^3c(2c-1)\right\}}{(c+x)(1+cx)(2c+x)(c+1+cx)(c+1)(2c+1)} \geq 0$$

and this proves the first part. For the second part we wish to show that the difference between (7.18) and (7.20) is non-negative and, after simplification, this reduces to showing that for $x \geq 1$ and $0 < c \leq \frac{1}{2}$

$$(7.23)\qquad x^2(1-2c^3) + xc(3-4c-2c^2) - 2c \geq 0.$$

To show (7.23), it is sufficient to show for any c with $0 < c \leq \frac{1}{2}$ that the roots of the quadratic in x must be less than or equal to unity. We need only consider the larger of the two roots, which has a plus sign in front of the radical since $1-2c^3 > 0$ for $0 < c \leq \frac{1}{2}$. After simplification, the inequality to be shown takes the form

$$(1+c)\,(1+2c)\,(1-2c^3)\,(1-2c) \geq 0 \tag{7.24}$$

which is true for $0 < c \leq \frac{1}{2}$; this proves the second part.

In conclusion, it has been shown that if the specified P^* is between $\frac{1}{3}$ and $\frac{2}{3}$ then we set the right hand member of (7.19) $\left[\text{worst case is } \sigma^2_{[1]} = \sigma^2_{[2]} = \sigma^2_{[3]}\right]$ equal to P^* and solve the resulting quadratic, obtaining for the positive root the appropriate value of c,

$$c = \frac{3(1-P^*) + \sqrt{(1-P^*)\,(17-P^*)}}{4(1+P^*)} \tag{7.25}$$

If the specified P^* is between $\frac{2}{3}$ and 1 then we use (7.20) $\left[\text{worst case is } \sigma^2_{[1]} = 0,\ \sigma^2_{[2]} = \sigma^2_{[3]}\right]$ in a similar manner and obtain for the appropriate c-value the simple result,

$$c = \frac{1-P^*}{P^*} \tag{7.26}$$

It is of some interest to note that if (7.25) is incorrectly used for P^* between $\frac{2}{3}$ and 1 then the maximum error attained in the guaranteed value of P^* is easily shown to be $7-4\sqrt{3} = .07$ to two decimal places.

REFERENCES

[1] Bechhofer, R. E., "A Single-Sample Multiple Decision Procedure for Ranking Means of Normal Populations with Known Variances," Ann. Math. Stat., Vol. 25 (1954), pp. 16-29.

[2] Bechhofer, R. E. and Sobel, M., "A Single-Sample Multiple Decision Procedure for Ranking Variances of Normal Populations," Ann. Math. Stat. Vol 25 (1954), pp. 273-289.

[3] Gupta S. S., "On a Decision Rule for a Problem in Ranking Means," Mimeograph Series No. 150, May, 1956, Institute of Statistics, University of North Carolina, Chapel Hill, N.C.

[4] Gupta, S. S. and Sobel, M., "On a Statistic Which Arises in Selection and Ranking Problems," Ann. Math. Stat. Vol. (28), 1957, pp. 957-967.

[5] Gupta, S. S. and Sobel, M., "Selecting a Subset Containing the Best of Several Binomial Populations," Chapter XX, pp. 224-248 in Contributions to Probability and Statistics, Stanford University Press, 1960.

[6] Gupta, S. S. and Sobel, M., "The Distribution and Percentage Points of the Smallest of Several Correlated F Statistics," to be submitted for publication.

[7] National Bureau of Standards, Tables computed for R. E. Bechhofer sponsored by ONR and distributed by the Chemical Corps. Engineering Agency, Army Chemical Center, Maryland, in the pamphlet, "Probabilities Associated with Order Statistics in Samples from Two Normal Populations with Equal Variance," by D. Teichroew (1955).